ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Journal of Number Theory 128 (2008) 2823–2835

JOURNAL OF Number Theory

www.elsevier.com/locate/jnt

Cohen–Lenstra heuristic and roots of unity

Gunter Malle

FB Mathematik, Universität Kaiserslautern, Postfach 3049, D-67653 Kaiserslautern, Germany

Received 19 November 2007

Available online 18 April 2008

Communicated by M. Pohst

Abstract

We report on computational results indicating that the well-known Cohen–Lenstra–Martinet heuristic for class groups of number fields may fail in many situations. In particular, the underlying assumption that the frequency of groups is governed essentially by the reciprocal of the order of their automorphism groups, does not seem to be valid in those cases. The phenomenon is related to the presence of roots of unity in the base field or in intermediate fields. For all the examples considered, we propose alternative predictions which agree closely with the data, and which are inspired by results of Gerth.

MSC: primary 11R29; secondary 11R16

1. Introduction

The multiplicative structure of a number field K is controlled by the class group Cl_K of its ring of integers O_K. Although efficient algorithms have been developed to compute the class group for any given number field K, the knowledge on the distribution of class groups is extremely weak. In 1983, Cohen and Lenstra [5] proposed a heuristic which makes predictions for various question of an asymptotic nature on class groups of quadratic number fields, and more generally abelian extensions of $\mathbb{Q}$. This was subsequently extended by Cohen and Martinet [7] to arbitrary number fields, see also [6] for tables of predictions in small degree.

In this paper we present numerical evidence indicating that this heuristic is not applicable to the p-part of the class group in the case where the base field or some intermediate field contains pth roots of unity. In particular, it seems to fail always when $p = 2$. Secondly, in some cases

E-mail address: malle@mathematik.uni-kl.de.

0022-314X/$ – see front matter
doi:10.1016/j.jnt.2008.01.002

we modify the original formulas to make new predictions which agree quite closely with our computational results.

Consider a *situation* $\Sigma := (K_0, G, \sigma)$ consisting of a number field K_0, a transitive permutation group G of degree $n \geqslant 2$, and a possible signature σ of a degree n extension K/K_0 with Galois group (of the Galois closure) permutation isomorphic to G. For such a situation Σ, let $\mathcal{K}(\Sigma)$ denote the degree n extensions K/K_0 of K_0 (inside a fixed algebraic closure) with group G and signature σ. We are interested in the structure of the relative class group of K/K_0 for $K \in \mathcal{K}(\Sigma)$ (the kernel in C_K of the norm map from K to K_0). The original heuristic of Cohen and Lenstra can now be phrased as follows: There exists a notion of primes which are *good for* Σ such that:

> For any prime p which is good for Σ the distribution of Sylow p-subgroups of class groups for fields in $\mathcal{K}(\Sigma)$ is described by a simple law which essentially weighs groups by the inverse of the order of their automorphism group.

Despite partial results, notably by Davenport–Heilbronn and Bhargava, there is no situation in which this prediction has been proved completely for even a single prime. Hence, no example of a good prime is known. Nevertheless, there are several proposals for what good primes should be: Clearly, by genus theory primes which divide the degree n cannot be good. Thus at most the primes in

$$\mathcal{P}_1(\Sigma) := \{p \text{ prime} \mid \gcd(p, n) = 1\}$$

can be good. Cohen, Lenstra and Martinet propose

$$\mathcal{P}_2(\Sigma) := \{p \text{ prime} \mid \gcd(p, |G|) = 1\} \subseteq \mathcal{P}_1(\Sigma)$$

as candidates for good primes.

All computational data obtained previously seemed in accordance with the conjecture that all primes in $\mathcal{P}_1$ are good. But, in fact, extensive computations have really only been carried out for quadratic number fields.

Here we present numerical data for the following situations Σ and primes p:

(1) $\Sigma = (\mathbb{Q}, \mathfrak{S}_3, \text{totally real})$, $p = 2 \in \mathcal{P}_1(\Sigma)$,
(2) $\Sigma = (\mathbb{Q}, \mathfrak{S}_3, \text{complex})$, $p = 2 \in \mathcal{P}_1(\Sigma)$,
(3) $\Sigma = (\mathbb{Q}(\sqrt{-3}), C_2, \text{complex})$, $p = 3 \in \mathcal{P}_2(\Sigma)$,
(4) $\Sigma = (\mathbb{Q}, C_3, \text{totally real})$, $p = 2 \in \mathcal{P}_2(\Sigma)$,
(5) $\Sigma = (\mathbb{Q}, C_5, \text{totally real})$, $p = 2 \in \mathcal{P}_2(\Sigma)$.

In all these situations our experimental data suggest that the respective p is *not* good for Σ. On the other hand, all other primes in $\mathcal{P}_1(\Sigma)$ behave as expected. Note that in all cases, the listed primes are such that K_0 contains the pth roots of unity. In view of this it seems that at most primes in

$$\mathcal{P}_3(\Sigma) := \{p \text{ prime} \mid \gcd(p, n) = 1 \text{ and } \mu_p \not\subseteq K_0\}$$

can be good, where μ_p denotes the pth roots of unity. In particular, the prime 2 would never be good. Furthermore, for extensions of composite degree, or more precisely, for non-primitive

Galois groups, even less primes should be good. For example, we will argue in Section 4 why the prime 3 should be bad for quartic dihedral extensions of $\mathbb{Q}$.

It was probably first noticed by Gerth [9] in cases where $p \notin \mathcal{P}_1$ (which are certainly bad primes) that the roots of unity contained in the base field influence the distribution of class groups. He showed that the 4-rank of class groups of quadratic extensions K/K_0 of imaginary quadratic fields K_0 depends on whether K_0 contains the fourth roots of unity, that is, whether $K_0 = \mathbb{Q}(\sqrt{-1})$ or not. (See also his paper [10] which is concerned with cyclic extensions of degree p of certain fields containing the pth roots of unity.) This did not conflict with the original Cohen–Lenstra heuristic. Nevertheless, Gerth had extended the original heuristic to include the prime 2, by assuming that the only difference to the case of odd primes comes from genus theory. (A similar extension to cyclic fields of prime degree was given by Gerth and Wittmann.)

Our computations originated in the investigation of totally real quartic $\mathfrak{A}_4$-extensions of $\mathbb{Q}$. Elements of order 2 in the class group of a C_3-field give rise to $\mathfrak{A}_4$-fields unramified over the C_3-subfield. Thus, the 2-rank of the class group of C_3-fields has an intimate relation to the asymptotic of the counting function for $\mathfrak{A}_4$-fields over $\mathbb{Q}$. Any attempt to understand this latter counting function must, in our opinion, also involve some result on the former.

2. Cyclic cubic fields

We first present computational results on class groups of cyclic cubic fields, that is, Galois extensions $K/\mathbb{Q}$ of degree 3. These computations are feasible since there exists an efficient algorithm for the generation of cyclic cubic number fields of bounded discriminant, almost linear in the number of fields, described for example in [3, 6.4.2].

We have obtained several sets of data. The first consists of all C_3-fields of discriminant at most 10^{16}; there are roughly 16 million such fields. Already, these fields together with discriminants and class groups occupy almost 900 MB of disc space. It hence did not seem reasonable to compute complete tables for even larger discriminants. Instead, we have also computed the first 100,000 fields of discriminant at least 10^i, for $i \in \{16, 18, 20, 22, 24\}$. The third set of data consists of the first 2,000,000 C_3-fields of prime conductor of discriminant at least 10^i, for $i \in \{14, 16, 18, 20\}$. Finally, we computed the first 6,000,000 C_3-fields of prime conductor of discriminant at least 10^{22}, and the first 1,000,000 such fields of discriminant at least 10^{28}.

For all these fields, we calculated the class group using the computer algebra system PARI/GP (http://pari.math.u-bordeaux.fr). This took roughly 3 months of computing time on a SUN workstation. See Section 2.4 for remarks on the reliability of the data.

Some of the results are displayed in the following tables, together with the corresponding predictions by Cohen, Lenstra and Martinet [5–7].

2.1. The odd behaviour of the even prime

In Table 1 we give the proportion of C_3-fields of discriminant bounded by $D = 10^{2i}$, $4 \leqslant i \leqslant 8$, with $3'$-part $h'(K) := |\mathrm{Cl}_K|_{3'}$ of the class number equal to h'. The last line of the table gives the proportions predicted by Cohen and Martinet in [6, (2)(a)].

It can be observed that the results for odd h' are as expected, but that even class numbers occur far too often.

Table 1
Proportion of h' for C_3-fields

D	h'									
	1	4	7	13	16	19	25	28	31	37
10^8	.837	.100	.031	.0069	.0050	.00691	.00188	.00440	.00188	.00063
10^{10}	.819	.108	.031	.0096	.0066	.00466	.00157	.00397	.00132	.00138
10^{12}	.804	.113	.034	.0096	.0095	.00461	.00148	.00460	.00160	.00137
10^{14}	.797	.117	.036	.0099	.0103	.00460	.00143	.00527	.00166	.00119
10^{16}	.794	.119	.037	.0101	.0110	.00462	.00140	.00557	.00169	.00119
[6]	.850	.071	.040	.0109	.0047	.00497	.00142	.00337	.00183	.00127

Table 2
Average elementary abelian p-part of class groups of C_3-fields

D	$\lvert S\rvert$	p					
		2	5	7	11	13	17
10^8	1 592	1.341	1.045	1.219	1.0000	1.083	1.0000
10^{10}	15 851	1.390	1.051	1.253	1.0076	1.144	1.0000
10^{12}	158 542	1.436	1.049	1.279	1.0136	1.156	1.0036
10^{14}	1 585 249	1.461	1.046	1.297	1.0096	1.161	1.0042
10^{16}	15 852 618	1.477	1.043	1.303	1.0090	1.161	1.0041
$\geqslant 10^{16}$	10^5	1.478	1.045	1.297	1.0024	1.165	1.0000
$\geqslant 10^{18}$	10^5	1.486	1.038	1.306	1.0096	1.164	1.0058
$\geqslant 10^{20}$	10^5	1.486	1.043	1.305	1.0084	1.151	1.0029
$\geqslant 10^{22}$	10^5	1.499	1.040	1.309	1.0120	1.167	1.0058
$\geqslant 10^{24}$	10^5	1.501	1.041	1.300	1.0096	1.160	1.0000
Prediction in [5]		1.250	1.040	1.306	1.0083	1.160	1.0035

We next consider the average elementary abelian p-part. For a prime p and a finite set S of cyclic cubic number fields let

$$\mathcal{M}\big(p^{\mathrm{rnk}_p}, S\big) := \frac{1}{|S|}\sum_{K\in S} p^{\mathrm{rnk}_p(\mathrm{Cl}_K)},$$

the average order of the maximal elementary abelian p-subgroup of the class group Cl_K, for $K \in S$. Table 2 gives $\mathcal{M}(p^{\mathrm{rnk}_p}, S)$ for some small primes $p \neq 3$ and for S of the form

$$\mathcal{K}(D) := \big\{K \;\big|\; \mathrm{Gal}(K/\mathbb{Q}) = C_3,\ d(K) \leqslant D\big\}$$

consisting of all cyclic cubic fields (inside a fixed algebraic closure of $\mathbb{Q}$) of discriminant bounded above by D, where $D = 10^{2i}$, $4 \leqslant i \leqslant 8$, as well as for S consisting of the first 100,000 C_3-fields of discriminant at least 10^{2i}, where $8 \leqslant i \leqslant 12$.

The heuristic in [5] predicts that

$$\lim_{D\to\infty} \mathcal{M}\big(p^{\mathrm{rnk}_p}, \mathcal{K}(D)\big) = \begin{cases} (1+\frac{1}{p})^2 & \text{for } p \equiv 1 \pmod 3,\\ 1+\frac{1}{p^2} & \text{for } p \equiv 2 \pmod 3,\end{cases}$$

corresponding to the splitting behaviour of p in the cyclotomic field $\mathbb{Q}(\zeta_3)$ of third roots of unity (see the last line of Table 2).

Again, the table shows two things. First, for odd primes $p \geqslant 5$, the actual data come rather close to the predictions, with increasing precision for larger D. Secondly, for $p = 2$, the data behave differently. In fact, the tables suggest that $\mathcal{M}(2^{\mathrm{rnk}_2}, \mathcal{K}(D)) \to 1 + \frac{1}{2}$ for $D \to \infty$, instead of $1 + \frac{1}{4}$.

2.2. *The* 2*-rank*

We next have a closer look at the behaviour of the 2-part of the class groups and illustrate the unexpected behaviour with some further statistics from our computations. It is predicted in [5, Ex. 5.9 and (C13)] and [6, 2(c)] that the probability $\mathrm{pr}(\mathrm{rnk}_2(\mathrm{Cl}_K) = r)$ that the class group Cl_K has a given (necessarily even) 2-rank r should tend to

$$2^{-r(r+2)/2}\frac{(4)_\infty}{(4)_{r/2}(4)_{r/2+1}} = \frac{(4)_\infty}{4^{r/2(r/2+1)}(4)_{r/2}(4)_{r/2+1}},$$

where

$$(q)_k := \prod_{i=1}^{k}\bigl(1 - q^{-i}\bigr) \quad \text{and} \quad (q)_\infty := \prod_{i=1}^{\infty}\bigl(1 - q^{-i}\bigr)$$

for $q \in \mathbb{N}$. Furthermore, the higher moments

$$\mathcal{M}\bigl(2^{n\,\mathrm{rnk}_2}, S\bigr) := \frac{1}{|S|}\sum_{K\in S} 2^{n\,\mathrm{rnk}_2(\mathrm{Cl}_K)} \quad (n \in \mathbb{N})$$

are predicted in [6, 2(d)] to tend to

$$\lim_{D\to\infty} \mathcal{M}\bigl(2^{n\,\mathrm{rnk}_2}, \mathcal{K}(D)\bigr) = \sum_{i=0}^{n} 4^{i(n-i-1)}\frac{(4)_n}{(4)_i(4)_{n-i}}.$$

Inspired by the result of Gerth [9, Theorem 1.1(ii)] on 4-ranks of class groups of quadratic extensions of $\mathbb{Q}(\sqrt{-1})$ we instead propose the following limit for 2-ranks:

$$\mathrm{pr}\bigl(\mathrm{rnk}_2(\mathrm{Cl}_K) = r\bigr) = \frac{3}{2}\cdot\frac{(2)_\infty(16)_\infty}{(4)_\infty^2}\cdot\frac{1}{4^{r/2(r/2+2)/2}(4)_{r/2}}. \tag{1}$$

This is indeed a probability measure:

Lemma 2.1. *We have*

$$\sum_{j=0}^{\infty}\frac{1}{4^{j(j+2)/2}(4)_j} = \frac{2}{3}\frac{(4)_\infty^2}{(2)_\infty(16)_\infty}.$$

Proof. The left-hand side equals

$$\sum_{j=0}^{\infty} \frac{4^{-j(j+2)/2}}{(4)_j} = \sum_{j=0}^{\infty} \frac{4^{-j(j-1)/2} 4^{-3j/2}}{(4)_j} = \prod_{k=0}^{\infty} (1 + 4^{-3/2} 4^{-k})$$

by [1, Corollary 2.2]. Now

$$\begin{aligned}\prod_{k=0}^{\infty} (1 + 4^{-3/2} 4^{-k}) &= \prod_{k=0}^{\infty} \frac{(1 + 2^{-3-2k})(1 + 2^{-3-2k+1})}{(1 + 2^{-3-2k+1})} \\ &= \frac{\prod_{k=3}^{\infty} (1 + 2^{-k})}{\prod_{k=2}^{\infty} (1 + 4^{-k})} = \frac{\prod_{k=3}^{\infty} (1 - 4^{-k}) \prod_{k=2}^{\infty} (1 - 4^{-k})}{\prod_{k=3}^{\infty} (1 - 2^{-k}) \prod_{k=2}^{\infty} (1 - 16^{-k})} \\ &= \frac{2}{3} \frac{(4)_\infty^2}{(2)_\infty (16)_\infty}\end{aligned}$$

as claimed. □

With this, the higher moments are given as follows:

Lemma 2.2. *The nth higher moment for the* 2*-ranks distributed as in* (1) *equals*

$$\prod_{k=1}^{n} (1 + 2^{2k-3}).$$

Proof. As in the previous proof we have

$$\sum_{j=0}^{\infty} \frac{4^{-j(j+2)/2}}{(4)_j} 2^{2nj} = \sum_{j=0}^{\infty} \frac{4^{-j(j-1)/2} 4^{j(n-3/2)}}{(4)_j} = \prod_{k=0}^{\infty} (1 + 4^{n-3/2} 4^{-k})$$

by [1, Corollary 2.2]. The result follows. □

In Table 3 we tabulate the average elementary abelian 2-part of the class group and its first few higher moments for fields K in the range of $d(K) \leqslant 10^{16}$, as well as for the first 100,000 fields above 10^{2i}, $9 \leqslant i \leqslant 12$, and compare this with the prediction from [5,6] and with our new formula (1).

It can be observed that the actual 2-ranks tend to be considerably larger than predicted in [6]. In fact, the larger the rank, the stronger the deviation becomes. For example, the original heuristic predicts less than 2 fields of 2-rank at least 6 in our range, while there are actually 243. The same holds for the higher moments.

On the other hand, the values come rather close to those predicted by formula (1).

In order to rule out that the behaviour of 2-parts is caused by special properties of some small primes, which then occur in many of the first composite discriminants, we have also compiled tables for the 2-part of class groups of C_3-fields of prime conductor. Here, in Table 4 the sets S consist of the first 2,000,000 fields of prime conductor and discriminant at least 10^i, where

Table 3
2-ranks and higher moments of class groups of C_3-fields

D	$\lvert S\rvert$	r				n			
		0	2	4	6	1	2	3	4
$\leqslant 10^{16}$	15 852 618	.8574	.1385	.00404	.15E–4	1.4771	4.169	30.14	548.0
$\geqslant 10^{18}$	10^5	.8553	.1403	.00429	.10E–4	1.486	4.24	30.0	485.7
$\geqslant 10^{20}$	10^5	.8553	.1405	.00421	.30E–4	1.486	4.30	35.0	816.0
$\geqslant 10^{22}$	10^5	.8532	.1421	.00472	.30E–4	1.499	4.46	37.1	849.9
$\geqslant 10^{24}$	10^5	.8538	.1412	.00501	.40E–4	1.501	4.56	40.9	1036.4
Formula (1)		.8530	.1422	.00474	.38E–4	1.500	4.50	40.5	1336.5
CL-prediction		.9180	.0816	.00035	.86E–7	1.250	2.31	7.6	45.9

Table 4
C_3-fields of prime conductor: 2-ranks and higher moments

D	$\lvert S\rvert$	r				n			
		0	2	4	6	1	2	3	4
$\geqslant 10^{18}$	$2 \cdot 10^6$	.8541	.1413	.00453	.29E–4	1.494	4.393	36.04	820.2
$\geqslant 10^{20}$	$2 \cdot 10^6$	.8541	.1412	.00465	.31E–4	1.495	4.431	37.06	861.8
$\geqslant 10^{22}$	$6 \cdot 10^6$	.8535	.1417	.00469	.41E–4	1.498	4.491	39.93	1035.2
$\geqslant 10^{28}$	10^6	.8527	.1426	.00464	.44E–4	1.500	4.501	40.44	1074.8
Formula (1)		.8530	.1422	.00474	.38E–4	1.500	4.500	40.50	1336.5
CL-prediction		.9180	.0816	.00035	.86E–7	1.250	2.312	7.58	45.9

$i \in \{18, 20\}$, as well as the first 6,000,000 fields of prime conductor and discriminant at least 10^{22}, and the first 1,000,000 fields of prime conductor and discriminant at least 10^{28}.

The same pattern as in the previous tables persists. In particular, as expected, the distribution of the 2-part of the class group is not influenced by the number of prime factors of the discriminant.

2.3. *The Sylow* 2*-subgroups*

We refine the previous tables by listing, in Table 5, the relative proportions of certain 2-groups as Sylow 2-subgroups of class groups of C_3-fields. Only those 2-groups which occur more than 200 times in our range have been retained. Also, in Table 6 we give the corresponding results for fields of prime conductor.

Let $A = \mathbb{Z}[\zeta_3]$. It is predicted in [6] that a given 2-torsion A-module H occurs with relative frequency

$$\frac{(4)_\infty}{(4)_1} \cdot \frac{1}{|H| \cdot |\mathrm{Aut}_A(H)|}$$

as Sylow 2-subgroup of a class group. Moreover, [7, Theorem 2.11] gives an explicit formula for $|\mathrm{Aut}_A(H)|$. The predicted values for some small 2-groups are given in the last line of Tables 5 and 6 respectively (see also [6, 2(a)]).

Once again, the observed frequency of all non-trivial 2-groups is considerably larger than was predicted, with almost identical behaviour for prime and non-prime conductor. Even more seri-

Table 5
Sylow 2-subgroups of class groups of C_3-fields

D	1	2^2	4^2	2^4	8^2	$4^2 \times 2^2$	2^6	16^2	$8^2 \times 2^2$
$\leqslant 10^{16}$	.857	.130	.0083	.00370	.00053	.32E–3	.13E–4	.38E–4	.21E–4
$\geqslant 10^{18}$	.855	.132	.0078	.00394	.00057	.29E–3	.10E–4	.10E–4	.50E–4
$\geqslant 10^{20}$	.855	.132	.0084	.00384	.00039	.34E–3	.20E–4	.40E–4	.20E–4
$\geqslant 10^{22}$	.853	.133	.0083	.00424	.00052	.44E–3	.30E–4	.20E–4	.40E–4
$\geqslant 10^{24}$	.854	.132	.0089	.00465	.00039	.36E–3	.40E–4	.10E–4	0
[6]	.918	.076	.0048	.00032	.00030	.25E–4	.79E–7	.19E–4	.16E–5

Table 6
C_3-fields of prime conductor: Sylow 2-subgroups

D	1	2^2	4^2	2^4	8^2	$4^2 \times 2^2$	2^6	16^2	$8^2 \times 2^2$
$\geqslant 10^{18}$	.854	.132	.0084	.00414	.00051	.36E–3	.27E–4	.35E–4	.25E–4
$\geqslant 10^{20}$	.854	.133	.0082	.00427	.00049	.35E–3	.30E–4	.35E–4	.19E–4
$\geqslant 10^{22}$	.854	.133	.0083	.00430	.00053	.34E–3	.43E–4	.31E–4	.25E–4
$\geqslant 10^{28}$	.853	.134	.0084	.00429	.00055	.32E–3	.40E–4	.41E–4	.28E–4
[6]	.918	.076	.0048	.00032	.00030	.25E–4	.79E–7	.19E–4	.16E–5

Table 7
Sylow 2-subgroups: deviation from [6]

H_2		H_2		H_2	
2^2	1.86	2^4	14.1	2^6	402.7
4^2	1.86	$4^2 \times 2^2$	15.0		
8^2	1.85	$8^2 \times 2^2$	15.3		
16^2	1.91	$16^2 \times 2^2$	11.0		
32^2	1.73	4^4	22.4		
64^2	1.84				

ously, the proportion among Sylow subgroups of a fixed order seems wrong. The fundamental assumption in [5], which also underlies the subsequent papers [6,7], is that among groups of the same order, the relative frequency should be inversely proportional to the order of the automorphism group. The smallest example here is given by the two groups of order 16, viz. C_4^2 and C_2^4. Here the A-automorphism groups have order 12 respectively 180, thus C_4^2 should occur about $180/12 = 15$ times as often as C_2^4 (see [6, (2)(a)]). The table shows that in the range of our data, this factor hovers around 2 instead. The situation is even worse for groups of order 64, but there the number of cases is still relatively small.

In Table 7 we have tabulated the medium deviation from the predictions [6,7] taken over all fields of prime conductor in Table 6 of discriminant at least 10^{16}. Here, the first column indicates the group structure of the Sylow 2-subgroup H_2 of Cl_K, the second gives the observed mean deviation.

We can make the following striking observation from this table: the deviation from the Cohen–Lenstra prediction essentially only depends on the *rank* of the Sylow 2-subgroup H_2. The higher the rank of H_2, the larger the factor, almost independent of the precise structure (and hence of

the automorphism group) of H_2. (There are just 25 fields with Sylow 2-subgroup of type C_4^4 and 17 fields with group $C_{16}^2 \times C_2^2$ in our range; therefor the corresponding entries should be taken cum grano salis.)

2.4. Reliability of the data

Our number of C_3-fields of discriminant at most 10^{16} agrees with the number published by Cohen, Diaz y Diaz and Olivier. For the fields of discriminant above 10^{16}, the numbers are in accordance with the proven asymptotic for C_3-fields.

The class groups were computed with the command `bnfclgp` in PARI/GP, version 2.4.1. The correctness of the output of this program relies on the (yet unproven) generalized Riemann hypothesis. For reasons of computing time, it is not feasible to check all class groups only using proven bounds. Thus the tables of class groups are not mathematically proven; on the other hand, any error would imply failure of GRH.

In order to confirm the deviation from the Cohen–Lenstra heuristic observed above, it suffices to consider those fields for which PARI/GP predicts an even class number. There are 2,260,230 such fields of discriminant at most 10^{16}. For all those fields of discriminant at most 10^{15}, we constructed all corresponding $\mathfrak{A}_4$-fields unramified over the intermediate C_3-field, which proves that the 2-rank of the class group in these cases is at least as large as predicted.

We have also checked all class groups with 2-rank at least 6 with the command `OrderClassGroup` in KANT (http://www.math.tu-berlin.de/~kant/kash.html) (which uses proven bounds) and constructed the corresponding unramified extensions, without finding a discrepancy.

3. C_5- and C_7-fields

It is now tempting to see whether the exceptional behaviour of the prime 2 extends to other cyclic extensions of prime degree. We have made some less extensive computations for cyclic extensions of $\mathbb{Q}$ of degree 5. Here, it is already quite time consuming to obtain a reasonable amount of data. Still, we managed to determine the class groups of the first 660,000 C_5-fields. The C_5-fields were constructed with the PARI/GP-command `rnfkummer`.

The 2-rank $\mathrm{rnk}_2(\mathrm{Cl}_K)$ of the class group of a C_5-field K is necessarily divisible by 4. By [5, Theorem 6.3 and Ex. 6.6] the probability $\mathrm{pr}(\mathrm{rnk}_2(\mathrm{Cl}_K) = r)$ of finding a given 2-rank $r \equiv 0 \pmod 4$ should tend to

$$\frac{(16)_\infty}{16^{r/4(r/4+1)}(16)_{r/4}(16)_{r/4+1}}$$

with higher moments

$$\sum_{i=0}^{n} 16^{i(n-i-1)} \frac{(16)_n}{(16)_i(16)_{n-i}}.$$

Instead, we propose that the 2-rank should be distributed by the analogue of our formula (1) for C_3-fields above, namely

$$\mathrm{pr}\big(\mathrm{rnk}_2(\mathrm{Cl}_K) = r\big) = \frac{5}{4} \cdot \frac{(4)_\infty (256)_\infty}{(16)_\infty^2} \cdot \frac{1}{16^{r/4(r/4+2)/2}(16)_{r/4}}. \tag{2}$$

Table 8
C_5-fields: 2-ranks and higher moments

D	$\lvert\mathcal{K}(D)\rvert$	r			n		
		0	4	8	1	2	3
10^{16}	647	.988	.0124	0	1.185	4.15	51.6
10^{20}	6 552	.986	.0136	0	1.204	4.46	56.6
10^{24}	65 634	.984	.0165	0	1.247	5.20	68.4
10^{28}	656 793	.984	.0162	.198E–4	1.248	6.43	399.4
Formula (2)		.984	.0164	.161E–4	1.250	6.25	406.2
CL-prediction		.996	.0042	.638E–7	1.062	2.07	19.1

Again, as in Lemma 2.1 it can easily be checked that this defines a probability measure, and as in Lemma 2.2 the corresponding higher moments are then given by

$$\prod_{k=1}^{n}\bigl(1+2^{4k-6}\bigr).$$

The computational data together with the Cohen–Lenstra prediction and with our new prediction (2) are collected in Table 8.

The behaviour of the 2-part of the class group is again quite different from the original prediction (although one could certainly argue that here the amount of data is still too small to make some definite statement). On the other hand, the new prediction matches the data quite convincingly.

For C_7-fields, the computations are even more expensive, and we have only managed to compute the first 33,000 class groups. Again, the 2-parts are much larger experimentally than would be expected.

4. Quadratic extensions of $\mathbb{Q}(\sqrt{-3})$

As indicated in the introduction, we expect that the failure of the Cohen–Lenstra heuristic is related to the existence of pth roots of unity in the base field. Another small case of this type, which is still amenable to computations, occurs for quadratic extensions of the field $\mathbb{Q}(\sqrt{-3})$ of third roots of unity. Here, the 3-part of the class group should behave differently. Note that the prime 3 lies in $\mathcal{P}_2$, since it does not divide the degree of the Galois closure (even over $\mathbb{Q}$) of such an extension.

If 3 were a good prime, the distribution of 3-ranks r of class groups of such quadratic extensions $K/\mathbb{Q}(\sqrt{-3})$ should be given by the same formula as for real quadratic extension of $\mathbb{Q}$, namely

$$\frac{(3)_\infty}{3^{r(r+1)}(3)_r(3)_{r+1}},$$

with higher moments

$$\sum_{i=0}^{n} 3^{i(n-i-1)}\frac{(3)_n}{(3)_i(3)_{n-i}}.$$

Table 9
C_2-fields over $\mathbb{Q}(\sqrt{-3})$: 3-ranks and higher moments

D	$\lvert S\rvert$	r				n		
		0	1	2	3	1	2	3
$\geqslant 10^8$	10^6	.8677	.129	.0036	.09E–4	1.287	2.326	7.16
$\geqslant 10^{12}$	10^5	.8550	.140	.0052	.60E–4	1.323	2.577	9.59
$\geqslant 10^{16}$	$2\cdot 10^6$	.8528	.141	.0058	.71E–4	1.331	2.648	10.55
$\geqslant 10^{18}$	10^5	.8528	.141	.0058	.11E–3	1.332	2.674	11.05
$\geqslant 10^{20}$	10^6	.8520	.142	.0059	.75E–4	1.333	2.661	10.45
Formula (3)		.8520	.142	.0059	.76E–4	1.333	2.667	10.67
CL-prediction		.8402	.158	.0023	.33E–5	1.333	2.444	6.81

Instead we expect this probability to be given by

$$\operatorname{pr}\bigl(\operatorname{rnk}_3(\mathrm{Cl}_K)=r\bigr)=\frac{4}{3}\cdot\frac{(3)_\infty}{(9)_\infty}\cdot\frac{1}{3^{r(r+3)/2}(3)_r} \tag{3}$$

with higher moments

$$\prod_{k=1}^{n}\bigl(1+3^{k-2}\bigr).$$

(Exactly the same formula occurs in Gerth's paper [10, Theorem 2] for the 3-rank of class groups of certain cyclic extensions of degree 3 of $\mathbb{Q}(\sqrt{-3})$.) Note that both formulas give the same value $4/3$ for the first moment, which was proven to be correct by Davenport–Heilbronn.

It is pretty straightforward to enumerate quadratic extensions of $\mathbb{Q}(\sqrt{-3})$ of bounded discriminant. We obtain the numerical data in Table 9. Again, the new prediction fits the data better.

Now consider the situation $\Sigma=(\mathbb{Q}, D_4, \text{totally complex})$ of quartic D_4-extensions of $\mathbb{Q}$ containing a non-real quadratic subfield. The number of such quartic D_4-fields with discriminant bounded by x is known to grow linearly with x. On the other hand, it is easy to see that the number of quadratic extensions of $\mathbb{Q}(\sqrt{-3})$ with norm of the discriminant at most x also grows linearly with x. Furthermore, among these the ones with Galois group of the Galois closure over $\mathbb{Q}$ equal to D_4 has density 1. Thus, the quartic fields with $\mathbb{Q}(\sqrt{-3})$ as subfield have positive density among all fields in the situation Σ. We have just argued that the prime 3 is probably not good for quadratic extensions of $\mathbb{Q}(\sqrt{-3})$. Up to an interchange of two limits this conflicts with the assumption that 3 is good for Σ.

5. Non-Galois cubic fields

A further simple but interesting situation for the Cohen–Lenstra heuristic is given by non-Galois cubic extensions of $\mathbb{Q}$. Here, the prime 3 is certainly bad, but it is not clear what to expect about the prime $p=2$. After first assuming that 2 should be good, Cohen and Martinet themselves later speculated [8] that their heuristic might not apply for $p=2$, in the light of computational data which showed an extremely slow convergence, if at all, against the predictions. Quite recently, though, Bhargava [2, Theorem 5] proved that at least the first moment for the 2-rank, viz. $5/4$ for totally real extensions and $3/2$ for non-real extensions, is correct as predicted.

Table 10
Totally real $\mathfrak{S}_3$-fields: 2-ranks and higher moments

D	$\lvert S\rvert$	r				n			
		0	1	2	3	1	2	3	4
10^4	366	.973	.027	0	0	1.027	1.08	1.19	1.41
10^6	54 441	.898	.101	.0014	0	1.105	1.32	1.79	2.87
10^8	6 246 698	.845	.149	.0057	.03E–3	1.167	1.53	2.42	4.81
$\geqslant 10^8$	10^6	.836	.158	.0067	.05E–3	1.178	1.58	2.55	5.28
$\geqslant 10^9$	10^6	.822	.169	.0088	.11E–3	1.196	1.65	2.80	6.23
$\geqslant 10^{10}$	10^6	.811	.177	.0110	.19E–3	1.212	1.71	3.03	7.30
$\geqslant 10^{11}$	10^6	.804	.183	.0124	.27E–3	1.222	1.75	3.20	8.07
$\geqslant 10^{12}$	10^6	.798	.188	.0136	.37E–3	1.232	1.79	3.36	8.80
Formula (4)		.786	.197	.0164	.59E–3	1.250	1.87	3.75	11.25
CL-prediction		.770	.220	.0098	.09E–3	1.250	1.81	3.20	7.18

Table 11
Non-real $\mathfrak{S}_3$-fields: 2-ranks and higher moments

D	$\lvert S\rvert$	r				n			
		0	1	2	3	1	2	3	4
10^4	1 520	.842	.157	.0013	0	1.161	1.49	2.18	3.68
10^5	17 041	.785	.207	.0082	0	1.232	1.74	2.97	6.20
10^6	182 417	.735	.248	.0169	.006E–2	1.299	2.00	3.83	9.29
$\geqslant 10^7$	10^5	.692	.278	.0299	.062E–2	1.372	2.32	5.14	15.3
$\geqslant 10^8$	10^5	.671	.292	.0356	.122E–2	1.407	2.49	5.91	19.4
$\geqslant 10^9$	10^5	.660	.299	.0396	.169E–2	1.430	2.60	6.53	23.8
$\geqslant 10^{10}$	10^5	.649	.305	.0433	.229E–2	1.451	2.71	7.12	27.3
Formula (5)		.629	.315	.0524	.374E–2	1.500	3.00	9.00	45.0
CL-prediction		.578	.385	.0367	.070E–2	1.500	2.75	6.37	19.2

Using the efficient algorithm of Belabas for enumerating $\mathfrak{S}_3$-fields of bounded discriminant, see [4], we have computed analogous data as above for the 2-part of the class group in both signatures. These are given in Tables 10 and 11. Although the range of discriminants is rather small, and the data are not as clear as in the previous cases, we still propose that the original prediction does not apply: In the totally real case, instead of

$$\frac{(2)_\infty}{2^{r(r+2)}(2)_r(2)_{r+2}}$$

we should have

$$\operatorname{pr}\bigl(\operatorname{rnk}_2(\mathrm{Cl}_K) = r\bigr) = \frac{15}{8}\cdot\frac{(2)_\infty}{(4)_\infty}\cdot\frac{1}{2^{r(r+5)/2}(2)_r} \tag{4}$$

with higher moments

$$\prod_{k=1}^{n}\bigl(1+2^{k-3}\bigr).$$

In the non-real case, instead of

$$\frac{(2)_\infty}{2^{r(r+1)}(2)_r(2)_{r+1}}$$

we expect

$$\operatorname{pr}\bigl(\operatorname{rnk}_2(\mathrm{Cl}_K)=r\bigr)=\frac{3}{2}\cdot\frac{(2)_\infty}{(4)_\infty}\cdot\frac{1}{2^{r(r+3)/2}(2)_r} \tag{5}$$

with higher moments

$$\prod_{k=1}^{n}\bigl(1+2^{k-2}\bigr)$$

(formula (3) with 3 replaced by 2). Note that both (4) and (5) give the same first moment as the original prediction, which was proved by Bhargava.

To our knowledge, our computations extend substantially further than previous ones. Note however that the number of fields, and in particular the range of discriminants, is much smaller than for the C_3-fields and quadratic extensions of $\mathbb{Q}(\sqrt{-3})$ considered before.

Acknowledgments

I thank Hendrik W. Lenstra for the remark that roots of unity might matter, and Jürgen Klüners for helpful discussions.

References

[1] G.E. Andrews, The Theory of Partitions, Encyclopedia Math. Appl., vol. 2, Addison–Wesley, Reading, MA, 1976.
[2] M. Bhargava, The density of discriminants of quartic rings and fields, Ann. of Math. 162 (2005) 1031–1063.
[3] H. Cohen, A Course in Computational Algebraic Number Theory, Grad. Texts in Math., vol. 138, Springer-Verlag, Berlin, 1993.
[4] H. Cohen, Advanced Topics in Computational Number Theory, Grad. Texts in Math., vol. 193, Springer-Verlag, Berlin, 2000.
[5] H. Cohen, H.W. Lenstra Jr., Heuristics on class groups of number fields, in: Number Theory, Noordwijkerhout, 1983, in: Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 33–62.
[6] H. Cohen, J. Martinet, Class groups of number fields: Numerical heuristics, Math. Comp. 48 (1987) 123–137.
[7] H. Cohen, J. Martinet, Étude heuristique des groupes de classes des corps de nombres, J. Reine Angew. Math. 404 (1990) 39–76.
[8] H. Cohen, J. Martinet, Heuristics on class groups: Some good primes are not too good, Math. Comp. 63 (1994) 329–334.
[9] F. Gerth, The 4-class ranks of quadratic extensions of certain imaginary quadratic fields, Illinois J. Math. 33 (1989) 132–142.
[10] F. Gerth, On p-class groups of cyclic extensions of prime degree p of certain cyclotomic fields, Manuscripta Math. 70 (1990) 39–50.